

AD _____

GRANT NUMBER DAMD17-94-J-4371

TITLE: Computer Aided Breast Cancer Diagnosis

PRINCIPAL INVESTIGATOR: Carey E. Floyd, Ph.D.

CONTRACTING ORGANIZATION: Duke University Medical Center
Durham, North Carolina 27710

REPORT DATE: September 1997

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

DTIC QUALITY INSPECTED 1

19980617 062

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE September 1997	3. REPORT TYPE AND DATES COVERED Annual (23 Sep 96 - 22 Sep 97)		
4. TITLE AND SUBTITLE Computer Aided Breast Cancer Diagnosis		5. FUNDING NUMBERS DAMD17-94-J-4371		
6. AUTHOR(S) Carey E. Floyd, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Duke University Medical Center Durham, North Carolina 27710		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Commander U.S. Army Medical Research and Materiel Command Fort Detrick, Frederick, MD 21702-5012		10. SPONSORING/MONITORING AGENCY REPORT NUMBER		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited		12b. DISTRIBUTION CODE		
13. ABSTRACT (Maximum 200) The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from the patients' mammographic findings and medical history. In the third year of this project, we have acquired 200 new cases bring our total case database to over 700. We have further investigated an alternative network architecture for predicting malignancy: a genetic algorithm. A user interface was further developed for more efficient data entry and error checking was employed. The existing mammography data entry database was evaluated for use in the prediction system. These developments were targeted at the first specific aim of the grant: develop an artificial neural network to predict biopsy outcome from mammographic and history findings. In the fourth year, we will focus on the second specific aim: evaluate the improvement in radiologists' diagnostic performance when the computer diagnostic aid is provided. This implementation of an accurate CAD system will improve sensitivity, specificity, and consistency of breast cancer diagnosis and will provide a significant improvement in long term outcome for breast cancer patients.				
14. SUBJECT TERMS Mammography, Radiology, Computer-Aided Diagnosis, Neural Networks, Malignancy Prediction, Artificial Intelligence, Breast Cancer		15. NUMBER OF PAGES 12		
		16. PRICE CODE		
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

☒ Where copyrighted material is quoted, permission has been obtained to use such material.

☒ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

☒ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

☒ NA
In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

☒ NA
For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

☒ NA
In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

☒ NA
In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

☒ NA
In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Carey S. G. J. 27 Feb 98
PI - Signature Date

Table of Contents

Front Cover	1
Report Documentation Page	2
Foreword	3
Table of Contents	4
Introduction	5
Report on Research for 1997	8
Conclusions	11
Personnel	12

Introduction

The long range goal of this project is to improve the accuracy and consistency of breast cancer diagnosis by developing a Computer Aided Diagnosis (CAD) system for early prediction of breast cancer from patients' mammographic findings and medical history. Specifically, this system will predict the malignancy of non-palpable lesions that are examined with diagnostic mammography and are considered for biopsy. The goal is to improve the specificity of diagnosis with little loss of sensitivity thus significantly improving the positive predictive value of breast biopsy.

Toward this goal, we have developed an artificial neural network (ANN) to predict biopsy outcome from mammographic and history findings. In the first three years of the grant we have 1) developed a user interface for acquiring mammographic findings, 2) acquired 500 cases using the standardized BI-RADSTM reporting system, 3) trained and evaluated an ANN predictive model, 4) conducted a small prospective study, 5) examined the inter- and intra-observer variability of the reporting lexicon, 6) investigated reducing the number of active input features, and 7) examined the sensitivity of the system to the techniques used for sampling the data.

What follows is a point by point assessment of the progress for each task in the original statement of work:

Statement of Work

Task 1, Develop an ANN to predict biopsy outcome from mammographic and history findings.

Years 1-4

Development will start with the successful preliminary backpropagation network. The significant improvements needed include: 1) larger set of clinical cases to better represent the general patient population, 2) higher specificity while maintaining >98% sensitivity. The preliminary work will be extended as follows.

Year 1

- 1.1) Expand the number of input features, both mammographic and medical history. The ANN will be implemented on a workstation (SUN SPARC) to allow the size of the network to be enlarged. This will allow more medical history and radiological features to be included.

These tasks were all achieved in year one.

Year 2-4

- 1.2) Develop a time-series ANN to examine current as well as previous exams.

Note: this aim was dropped in response to the decreased budget as negotiated with BC Baker in a revised statement of work in August 1994 .

- 1.3) Evaluate other ANN architectures which have been demonstrated to be appropriate for pattern classification.

Achieved in year 2.

Year 3-4

- 2) Evaluate the improvement in radiologists' diagnostic performance when the computer diagnostic aid is provided.

Some of this task was achieved in year one.

Year 3

Install the trained network on the Mammography Database server to perform on-line prediction as the radiologists input the features.

Achieved in year 2. Documented below.

Year 3-4

Test the hypothesis that use of the network prediction by radiologists will increase diagnostic accuracy (prediction of biopsy results).

Not yet achieved. Begun in year 3.

In summary, we have achieved all work for year one, some of the work allocated to years 2-4, some of the work allocated to year 3, and some of the work allocated to years 3-4. We are on schedule and anticipate that we will complete all work by the end of year 4.

In the third year of the grant we have published five peer-reviewed manuscripts [1-5]. There have been 2 presentations with published proceedings at professional meetings [6, 7] and two abstracts published from international meetings [8, 9]. Specifically, we have acquired 200 new cases using the standardized BI-RADS reporting system bringing our total to 700 cases. All of this work has been specifically directed toward the first specific aim of the proposal.

In summary:

	Year 3:	Cumulative
Peer-reviewed manuscripts published or in press:	5	8
Published Conference Proceedings:	2	11
International Meeting presentations:	2	15
Related grants received:	2	4

Peer-reviewed manuscripts published or in press:

- 1 Baker JA, Kornguth PJ, Lo JY, Floyd CE Jr: An Artificial Neural Network Approach to Improve the Quality of Breast Biopsy Recommendations *Radiology*;198;131-135;1996.
- 2 Baker JA, Kornguth PK, Floyd CE Jr: BI-RADS Standardized Mammography Lexicon: Observer Variability of Lesion Description *Amer. J. Roent.* , Apr 1996.
3. Tourassi GD, Floyd CE Jr: The Effect of Data Sampling on the Performance Evaluation of Artificial Neural Networks in Medical Diagnosis. *Medical Decision Making* ;17;186-192; 1997.
4. Lo JY, Baker JA, Kornguth PJ, Igelhart R, Floyd CE Jr: Predicting Breast Cancer Invasion From BI-RADS Mammographic Features Using Artificial Neural Networks On The Basis Of Mammographic Features *Radiology* ;203;159-163;1997.
5. Floyd CE Jr, Lo JY, Tourassi GD, Baker JA, Vitittoe NF, Vargas-Voracek R: Computer Aided Diagnosis in Thoracic and Mammographic Radiology. *Medical Imaging Technology*;6;629-634;1996.

Published Conference Proceedings:

6. Floyd CE Jr: Use of genetic algorithms for computer-aided diagnosis of breast cancer from image features. In *Proceedings of the International Society for Optical Engineering (SPIE)*; 2710;51-58;1996.
7. Lo JY, Floyd CE Jr, Kornguth PJ. Computer-aided diagnosis of mammography using an artificial neural network: predicting the invasiveness of breast cancers from image features. In *Proceedings of the International Society for Optical Engineering (SPIE)*; 2710: 725-732; 1996.

Meeting presentations (in addition to those listed in the conference proceedings):

8. Lo JY, Baker JA, Kornguth PJ, Floyd CE Jr: Computer-Aided Diagnosis Of Mammography: Artificial Neural Networks For Optimized Merging Of Standardized BIRADS Features. Presented at *World Congress on Neural Networks*, International Neural Network Society Annual Meeting (INNS), 1996.
9. Lo JY, Baker JA, Floyd CE Jr: Artificial Neural Networks For The Prediction Of Breast Cancer Invasiveness By Using Breast Imaging And Reporting Data System Mammography Lexicon, *Radiology*;201P;370; 1996.

Report on Research for 1997

In the third year of the project, we continued the development of an artificial neural network (ANN) to assist radiologists in the differentiation of benign from malignant lesions. Inputs to the ANN were derived from the patient's history and the radiologist's description of lesion morphology following the ACR Breast Imaging Reporting and Data System (BI-RADSTM). The output of the neural network is the likelihood of malignancy.

Artificial neural networks are a form of artificial intelligence analogous to layers of biological neurons. These networks can be trained to "learn" essential information from a set of data. The structure of an ANN is a set of processing units (nodes) arranged in rows. Input nodes are interconnected by simple calculations with an internal layer of hidden nodes and a single output node. Rather than having a fixed algorithmic approach to a classification problem, an ANN is sequentially presented with a set of supervised training cases — input data paired with the correct output. The ANN modifies its behavior ("trains") by adjusting the strength or "weights" of the connections until its own output converges to the known correct output. The information "learned" by the ANN is stored in the weight the network gives to connections between nodes.

ORGANIZATION OF THE NEURAL NETWORK

The ANN for prediction of breast malignancy was constructed as a three layer feed-forward network with a backpropagation training algorithm. The layers consist of an input layer with 18 input nodes, one hidden layer with 10 nodes, and an output layer with one output node. Each input node corresponds to either a radiologist's description of a feature of the lesion or information from the patient's medical or family history.

Of those women undergoing needle localization for nonpalpable breast lesions, a total of 500 lesions were identified on these studies that went on to open excisional biopsy and pathological diagnosis.

Each set of mammograms was acquired using film-screen technique on dedicated mammography equipment. No case was included in the study if either of the reviewing radiologists had prior knowledge of the biopsy results or if the suspicious area was not definitely identified. Of the 500 lesions evaluated there were 232 masses alone, 192 suspicious calcifications, and 29 combinations of masses and associated microcalcifications. The remaining 47 lesions included various combinations of architectural distortion, regions of asymmetric breast density, areas of focal asymmetric density, and areas of asymmetric breast tissue. Patients ranged in age from 24 to 86 years with an average age of 55 years. At biopsy, 326 (65%) of the lesions were found to be benign while 174 (35%) were malignant. This PPV of 35% is somewhat greater than that described in prior studies.

Each set of training films was reviewed prospectively by one of two radiologists whose primary clinical responsibilities are the interpretation of mammograms and the evaluation of breast lesions and who are familiar with the definitions of the BI-RADSTM descriptors. At least two views of the breast with the suspicious lesion were provided to the participating radiologists; a cranio-caudal and mediolateral-oblique view were available in all cases. Other views including true lateral, magnification views, and spot compression views as well as comparisons with the opposite breast were provided for evaluation when available. In order to avoid biasing the radiologist's description of the lesion, films from prior studies and the patient's history were initially withheld while the reviewing radiologist chose descriptors for each lesion. The radiologist was asked to describe each lesion using the BI-RADSTM lexicon by completing a checklist that included all possible BI-RADSTM descriptors. The reviewing radiologist was permitted to select only a single descriptor from each category. Each reader was blinded to the biopsy results while reviewing the films. The lesion descriptors along with patient history were used as inputs to train a neural network as described below.

Finally, to compare the performance of the ANN to experienced radiologists, the reviewing mammographer was provided with the patient's history and any prior films to correlate with the study mammograms and was requested to estimate the likelihood of malignancy. A five point scale was used with 1=very likely benign, 2=likely benign, 3=indeterminate, 4=likely malignant, and 5=very likely malignant.

NETWORK INPUTS

A total of 18 inputs were used to train the ANN to distinguish benign from malignant lesions. Ten of the inputs consisted of morphologic features extracted from the lesion by a radiologist. The remaining 8 inputs encompassed data from the patient's personal and family history collected from a survey form completed by the patient at the time of the exam. Each input is information routinely collected using the ACR BI-RADSTM standardized lexicon.

The first three features are descriptive features that apply to microcalcifications and calcifications associated with masses: calcification distribution, number and description. Inputs four through seven apply only to masses: mass margin, mass shape, mass density, and mass size. Three descriptive features that can apply to all lesions include lesion location, associated findings (e.g. axillary adenopathy), and special cases (e.g. asymmetric breast tissue).

The remaining 8 inputs are data from each patient's history. These include the patient's age, history of prior breast cancer, history of prior ipsilateral benign biopsy, weak, intermediate or strong family history of breast cancer, menstrual status, and use of estrogen or progesterone therapy. All morphologic features and patient history data were assigned a numerical value which was then scaled so that each input

ranged from zero to one. The order of the inputs in each category was determined at the beginning of the study by discussion with experienced mammographers and review of reports discussing the malignant potential of various BI-RADSTM descriptors.

Table 1 Performance of the trained neural network

Performance: Sparing Benign Biopsies					
Sensitivity	Specificity	Positive Predictive Value	Malignancies Missed	Benign Biopsies Spared	ANN Output Threshold
100	0	35	0	0	0.000
100	22	41	0	72	0.025
98	41	47	4	133	0.081
95	52	51	9	168	0.119
90	64	57	17	208	0.175
85	69	59	26	225	0.216

This table shows the performance of the network as the decision threshold is varied.

Genetic algorithm

PURPOSE

In this investigation we have explored genetic algorithms as a technique to train the weights in a feed-forward neural network to predict breast cancer from mammographic findings and patient age. This is a continuation of work that was begun in the second year of the grant. The work was submitted for presentation at the SPIE meeting in Feb 1998.

Conclusions

No new difficulties have been identified. One difficulty was described last year. The original statement of work was based on the use of a computer database of mammographic findings. Since the time of submission, the clinical use of this database has changed. In addition, as we developed our data acquisition protocol, we found that some items that we needed were not available from the database. While we are negotiating the modification of the on-line data entry forms, we have been acquiring data using paper forms. These forms do not constitute much additional work for the mammographers and have been received with acceptance. We have acquired BI-RADS findings for every biopsy case for the last year. This paper-based data collection system is in place and we anticipate no interruption of data acquisition for the duration of the grant. Since the study section identified the on-line database as a strength of the grant proposal, we continue to actively work to straighten out the compromises required to achieve the on-line data collection. In truth, the difference between paper-based and on-line data collection has no effect on the scientific quality of the research. We have conducted a systematic comparison of the on-line database with the paper form database. For the first 500 cases, we found complete agreement for the most important findings: calcification description and mass margin. The analysis for the secondary findings is underway.

The performance of the current network is described in table 1. The system currently could avoid 72 of the 326 benign biopsies without missing a malignancy. With a less conservative approach, 40% of the benign biopsies could be avoided at the cost of missing 4 malignancies.

Personnel

Carey E. Floyd, Jr., PhD

Phyllis Kornguth, MD, PhD

Joseph Lo, PhD

Georgia Tourassi, PhD